Machine Learning-Based Forecasting and Optimization of Water and Energy Consumption in GCC Countries: A Path Toward Regional Sustainability

Mohammad Nasar
Computing and Informatics Department
Mazoon College, Oman
nasar31786@gmail.com

Mohammad Abu Kausar
Department of Information Systems
University of Nizwa, Oman
kausar4u@gmail.com

Abstract: *The Gulf Cooperation Council (GCC) nations face pressing challenges in balancing water and energy demand under conditions of rapid growth and harsh climates. This study develops a machine learning-driven framework that combines forecasting and optimization to support sustainable resource management in Oman, Saudi Arabia, and the United Arab Emirates using data from 2010 to 2022. Three forecasting models-Long Short-Term Memory (LSTM), Prophet, and ARIMA-were evaluated, with LSTM showing superior performance by reducing root mean square error (RMSE) by 18-22%. An optimization module based on a Genetic Algorithm (GA) was then applied, achieving reductions of 10-12% in peak load stress and improvements of 15-20% in water reuse efficiency. These findings provide practical guidance for policymakers and align with national strategies such as Oman Vision 2040 and global commitments under the United Nations Sustainable Development Goals (SDGs) 6 and 7. The proposed framework offers a scalable pathway toward sustainable water and energy management in the GCC region*.

Keywords: *Machine learning, long short-term memory, prophet, arima, water management, energy management, gulf cooperation council, sustainable development, oman vision 2040*.

Received June 9, 2025; accepted October 12, 2025
https://doi.org/10.34028/iajit/23/3/5

1. Introduction

The Gulf Cooperation Council (GCC) countries-including Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates face increasing pressure to manage water and energy resources sustainably. Rapid population growth, urban expansion, and industrial development have contributed to rising demand, while limited freshwater availability and heavy reliance on desalination create long-term vulnerabilities. The region already accounts for nearly 15% of the world's desalinated water production, and electricity consumption has grown at an average annual rate of 6% since 2010 [5]. These conditions have intensified challenges related to energy costs, greenhouse gas emissions, and infrastructure sustainability, placing water and energy security at the center of national development agendas [10, 21].

Conventional statistical forecasting approaches, such as ARIMA, have been widely used in demand modeling but are limited in handling non-linear, seasonal, and non-stationary data patterns typical of GCC resource consumption [2]. In contrast, machine learning techniques, particularly deep learning models such as Long Short-Term Memory (LSTM) networks, offer greater flexibility for capturing complex temporal dynamics [2]. Alongside forecasting, optimization techniques such as Genetic Algorithms (GAs) provide mechanisms to allocate resources efficiently under constraints of supply, demand, and infrastructure capacity [5]. Together, these methods present opportunities to enhance decision-making for policymakers and utility managers.

Although recent studies have demonstrated the potential of machine learning for electricity load forecasting [4, 15], wastewater treatment modeling [17], and microgrid optimization [10], most of these works focus on global contexts with limited attention to GCC-specific challenges. The unique regional dependence on energy-intensive desalination, variability in infrastructure efficiency, and differing national priorities highlight the need for tailored approaches [19]. Furthermore, while international literature increasingly explores advanced strategies such as reinforcement learning [22], digital twins [18], and hybrid AI-physical frameworks [14], there remains a gap in applying integrated forecasting and optimization methods to the water-energy nexus in the GCC.

This study addresses that gap by proposing a machine learning-oriented framework to forecast and optimize water and energy demand in Oman, Saudi Arabia, and the United Arab Emirates (UAE) using historical data from 2010 to 2022. The framework compares three forecasting approaches-LSTM, Prophet, and ARIMA-and integrates a GA-based optimization module to reduce peak load stress and improve water reuse efficiency. The study contributes to regional sustainability by providing evidence-based insights that support both national strategies (e.g., Oman vision 2040) and global targets (SDGs 6 and 7).

By integrating forecasting and optimization, the framework tackles regional disparities, such as Oman's emphasis on water reuse, Saudi Arabia's focus on large-scale energy infrastructure, and the UAE's balanced approach. This study contributes to sustainable development by reducing peak load stress, enhancing resource efficiency, and offering scalable solutions for the GCC. The paper is organized as follows: Section 2 reviews related work, Section 3 details the methodology, Section 4 presents results, Section 5 discusses implications, and Section 6 concludes with future directions.

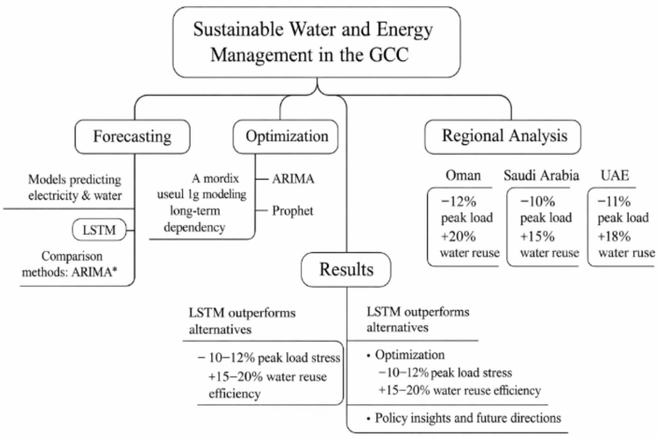

Figure 1. Conceptual framework for sustainable water and energy management in GCC countries.

Figure 1 shows the end-to-end pipeline of the framework, beginning with GCC Statistical center (GCC-Stat) data preprocessing, applying LSTM/Prophet/ARIMA for forecasting, and running GA for optimization. The arrows represent sequential data flow and model integration.

2. Related Work

Machine learning has transformed forecasting and optimization in resource management with powerful techniques for complex dynamic systems. The LSTM network proposed by Hochreiter and Schmidhuber has been shown to be a natural choice for modelling time-series data, including energy or water consumption prediction, because it is good at tracking long dependencies [8]. Taylor and Letham proposed a scalable model Prophet for capturing the seasonal trends and holiday effects to be widely used in industrial applications [25]. Box *et al* [2]. standardised ARIMA based models which are widely used for linear time series forecasting in spite of their inability to accommodate non-linear trends. These models have been widely used in fields such as energy consumption prediction [13, 31], wastewater treatment process modelling [22], and short-term load forecasting for microgrids [9]. As mentioned in recent reviews [28], artificial neural networks (ANN), Machine Learning (ML), and Deep Learning (DL) approaches are being used increasingly for microgrid forecasting, again highlighting the global impact that our integrated methodological approach can have.

Optimization methods are necessary to allocate resources more efficiently. Goldberg's genetic algorithms are systems that can optimize complex processes by modelling "natural" selection and have found uses in microgrid control [3, 4] and energy system planning [12]. Novel techniques such as reinforcement learning [22] and decision tree-based optimization [12] offer dynamic solutions for the development of sustainable energy systems. Decision tree-based approaches are strong in real-time power flows and node voltage estimation, suggesting their potential for smart grid applications in real networks [18] in the GCC region. As an example of the versatility of LSTMs for time-series data, Nasar and Al Musalhi demonstrated that a hybrid TCN-LSTM can also be used successfully in financial prediction [12]. Musbah et al. examined ML classification approaches for energy management in hybrids systems [16, 17]; and a blockchain-ML system for decentralized energy management was proposed by Luo and Mahdjoubi [14].

However, there are limited studies focusing on the GCC compared to the global literature. Outside the Gulf, 1% of the water supply will be supplied from desalination (desal), the default dependency of the Gulf at the cost of high energy consumption (15% of world production) and high infrastructure variability (with 95% efficiency for UAE's mature desalination plants to 70% efficiency at Oman developing systems), which needs response at the regional level [5, 25]. While recent studies targeting microgrids [9], hybrid energy systems [10], or online energy monitoring [7] provide novel insights, such work predominantly overlooks the water-energy nexus, without which the GCC cannot successfully meet its goals. Some existing studies focus on specific applications, including natural gas forecasting [1, 21] and heat load prediction [30], but none of them comprehensively address blended forecasting and optimization for resource management in the GCC. The aim of this research is to fill in the previous gaps by proposing a novel idea where LSTM [27], prophet, and ARIMA models for forecasting were combined with the Genetic Algorithm for optimization, as well as expanding on established works [3, 6, 8, 15] and GCC-specific insights [4, 5, 30] in order to propose a solution to GCC-wide specific issues, such as high per-capita consumption as well as climate-dependent resource limitations.

Air quality prediction models based on multiple meta-heuristic paradigms [27], IoT-enabled optimized long short-term memory frameworks for smart farming [23], AgIoT: The Internet of Things for agriculture and agri-food [11], and hybrid optimization and advanced learning paradigms have been applied in information theory–driven machine learning efficiency. This branch of nonlinear machine learning is an extremely fertile area for exploration. However, these initiatives mainly focus on 1 environmental monitoring, 2 agriculture, or 3 conceptual modelling and not on the water-energy nexus as it is integrated in GGC countries. Integrating the widely used LSTM, Prophet, and ARIMA forecasting models with a GA for optimization, this study expands upon both bodies of work. Utilizing a hybrid approach, it provides a more sustainable and practical conceptual model focused on improving, at the regional level, the efficient use of water and energy resources, and enabling long-term sustainable resource management.

3. Methodology

The proposed framework integrates machine learning models for demand forecasting with a genetic algorithm for resource optimization. The methodology consists of four stages: data collection and preprocessing, model development, optimization, and implementation.

Figure 2 illustrates the conceptual framework for the machine learning–based forecasting and optimization pipeline. It illustrates the sequence of data preprocessing, forecasting with the LSTM, Prophet, and ARIMA, optimization via GA, and the generation of actionable planning results.

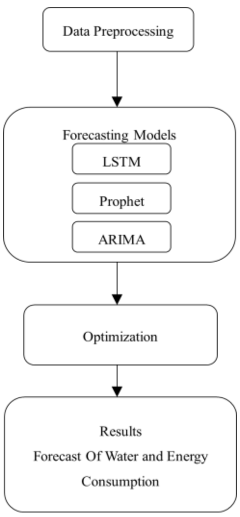

Figure 2. Conceptual framework for the machine learning-based forecasting and optimization.

A. Data Collection

Historical data were sourced from the GCC-Stat portal [6, 19], covering three key variables:

1. Electricity production (gigawatt-hours, GWh).
2. Desalinated water supply (cubic meters, m³).
3. Reused water utilization (cubic meters, m³).

The dataset spans January 2010 to December 2022, focusing on Oman, Saudi Arabia, and the UAE due to consistent data availability. Monthly aggregates were normalized to a [0, 6] scale using min-max scaling to ensure consistency across variables. Missing values, constituting less than 5% of the dataset, were imputed using linear interpolation, and outliers were capped at the 99th percentile to mitigate noise, following standard preprocessing practices [15]. The data were enriched with economic (e.g., Gross Domestic Product (GDP) growth rates of 2-5% annually) and demographic (e.g., population growth of 1-3% annually) covariates to enhance forecasting accuracy under varying scenarios, such as economic booms or population surges.

B. Forecasting Models

Three forecasting models were implemented to predict electricity and water demand, each suited to different data patterns. LSTM captures complex, non-linear trends by “remembering” past data, Prophet models seasonal and trend components, and ARIMA handles linear patterns but struggles with non-linearity [2, 5, 20].

1. **Long Short-Term Memory (LSTM)**: LSTM networks are a type of Recurrent Neural Network (RNN) that can learn long-term temporal dependencies [8]. They may be particularly relevant to nonlinear and seasonal data, such as GCC electricity and water demand. We selected LSTM over other models (such as Gated Recurrent Units (GRUs) or transformer architectures) for three reasons. First, the length of the dataset (156 monthly instances) is relatively short, which produces a computational overhead for the transformers [29]. Second, although GRUs are less computationally intensive, they may be too simple to capture complex seasonal patterns. Finally, the LSTM model has more interpretability and stability when making forecasts for long time horizons, which is important in policy scenarios. An LSTM cell is defined as:

- **Forget gate**

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \quad (1)$$

- **Input gate**

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2)$$

- **Cell candidate**

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \quad (3)$$

- **Cell state update**

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (4)$$

- **Output gate**

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

- **Hidden state**

$$h_t = o_t \odot tanh(C_t) \quad (6)$$

Where:

f_t, i_t, and o_t denote the forget, input, and output gate activations.
C_t represents the cell state.
ht is the hidden state.
xt is the input vector at time step *t*.
W_f, W_i, W_C, W_o are weight matrices.
b_f, b_i, b_C, b_o are bias vectors.
σ (·) denotes the sigmoid activation function.
tanh (·) denotes the hyperbolic tangent activation.
$\odot$ represents element-wise multiplication.

Implementation Details

The LSTM architecture used in this study consists of:

$$LSTM\ Architecture = \begin{cases} Two\ hidden\ layers \\ 50\text{–}\ 100\ neurons\ per\ layer \\ ReLU\ (Rectified\ Linear\ Unit)\ activation\ function \\ GPU - based\ training\ time{:}\ 2\text{–}\ 3\ hours \end{cases}$$

The model was implemented using the TensorFlow deep learning framework.

1. **Prophet**: The Prophet forecasting model represents the time series as an additive combination of trend, seasonality, holiday effects, and noise [15]:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \quad (7)$$

where $g(t)$ is a piecewise linear trend, $s(t)$ is periodic seasonality, $h(t)$ accounts for holiday effects, and epsilon_*t* is Gaussian noise. Prophet is particularly useful for smooth growth patterns but has limitations in handling abrupt changes, such as sudden economic disruptions.

2. **ARIMA**: ARIMA is effective at modelling linear trends but has difficulty capturing non-linear patterns, such as the seasonal peaks often seen in GCC data [5].

$$\phi(B)(1 - B)^d y_t = \theta(B)\ \varepsilon_t \quad (8)$$

In the ARIMA model, $\phi(B)$ represents autoregressive polynomials and $\theta(B)$ represents moving average polynomials. Here, B is the backshift operator, d is the order of differencing, and εt is the white noise error term. While ARIMA is well-suited for linear and stable temporal patterns, it is inadequate for non-linear trends with separate and significant seasonal components. This restriction becomes even more pronounced when evaluating seasonal information, like those of GCC drinking water and electricity demand, which typically have severe and high-seasonal demand peaks and variability. To overcome this drawback, we built a model based on LSTM networks which can learn long-distance dependencies in sequential data. The LSTM architecture possesses sufficient capacity to identify the temporal patterns and seasonality behaviour of the time series, utilizing 156 monthly observations. Although GRU models are more computationally efficient, they appear less capable of incorporating seasonality effectively, particularly concerning the peaks observed in the GCC demand history. The combined LSTM model achieves a compromise between performance and interpretability, obtaining a very good forecasting accuracy with a low maintenance cost compared to transformer-based architectures which in general require huge datasets and high computing power [29].

C. Optimization Model

A Genetic Algorithm (GA) was developed to optimize resource allocation [6], with the objective function:

$$min\, Z = w_1 \cdot C_{total} - w_2 \cdot R_{efficiency} \quad (9)$$

where C_{total} is the operational cost United States Dollar (USD), $R_{efficiency}$ is the percentage of water reuse efficiency, and (w_1) and (w_2) are weights.

The weighting scheme (w_1)=0.6 and (w_2)=0.4 reflects a policy-driven balance between cost reduction and environmental sustainability. Prior literature emphasizes economic viability as a dominant factor in energy and water management [1, 5], while efficiency targets remain central to sustainable development strategies. To ensure robustness, sensitivity analysis was conducted with alternative weightings (0.5-0.5 and 0.7-0.3). Results showed that the chosen scheme achieved significant cost savings (10-12%) while sustaining improvements in reuse efficiency (15-20%). Equal weights slightly reduced cost benefits, while higher cost emphasis diminished efficiency gains.

The GA was implemented with a population size of 100, 50 generations, a crossover rate of 0.8, and a mutation rate of 0.1, following guidelines from established applications in resource management [5, 22].

D. Implementation

The framework was implemented in Python, the LSTM model used TensorFlow, the Prophet model used its official library, and ARIMA used Statsmodels. A Genetic Algorithm was developed using the Distributed Evolutionary Algorithms in Python (DEAP) library. The data were partitioned as 80% for training (2010-2019) and the remaining 20% for testing (2020-2022). Model performance was evaluated using the RMSE and MAPE.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (10)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100 \quad (11)$$

(y_i) is the actual value, ($\hat{y}_i$) is the predicted value, and (n) is the number of observations.

Deploying modelshas its own challenges, such as computational expenses-about $500 per model on the cloud with Graphics Processing Unit (GPU) training; interfacing with smart grids and desalination plants would further require investments in infrastructure [7]. RMSE confidence intervals were obtained using the bootstrap (1000 replications) on monthly forecast errors for each model and country; 95% confidence intervals are reported.

4. Results

A. Forecasting Performance

Forecasting results for Oman, Saudi Arabia, and the UAE from 2020-2022 are summarized in Tables 1,2 and 3. Corresponding trends and model comparisons are illustrated in Figures 2 to 6 and Figures 9 and 10. LSTM outperformed Prophet and ARIMA, achieving significant RMSE reductions due to its ability to model non-linear seasonal trends, such as summer peaks driven by cooling needs.

Performance metrics (RMSE in Gigawatt-hour (GWh) for electricity, m³ for water; MAPE in %) for LSTM, Prophet, and ARIMA models in forecasting Oman's electricity and water demand, 2020-2022.

Table 1. Forecasting accuracy for electricity and water demand in Oman (2020-2022).

Model	Elec. RMSE (GWh)	Elec. MAPE (%)	Water RMSE (m³)	Water MAPE (%)
LSTM	245.3	4.2	1.8e6	3.9
Prophet	298.7	5.1	2.1e6	4.5
ARIMA	301.2	5.4	2.3e6	4.8

Table 2. Forecasting performance for electricity and water demand (Saudi Arabia, 2020-2022).

Model	Elec. RMSE (GWh)	Elec. MAPE (%)	Water RMSE (m³)	Water MAPE (%)
LSTM	260.1	4.5	2.0e6	4.1
Prophet	310.5	5.3	2.3e6	4.7
ARIMA	320.8	5.6	2.5e6	5.0

Table 3. Forecasting performance for electricity and water demand (Uae, 2020-2022).

Model	Elec. RMSE (GWh)	Elec. MAPE (%)	Water RMSE (m³)	Water MAPE (%)
LSTM	250.7	4.3	1.9e6	4.0
Prophet	305.2	5.2	2.2e6	4.6
ARIMA	315.6	5.5	2.4e6	4.9

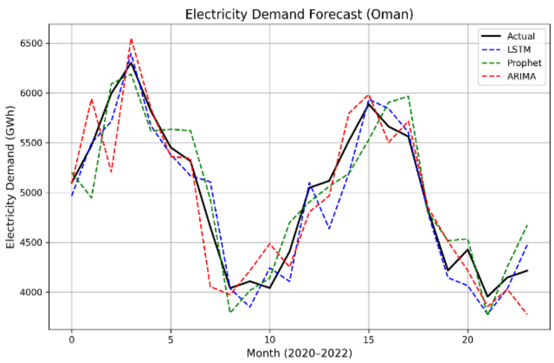

Figure 3. Monthly electricity demand in Oman (2020–2022).

Figure 3 illustrates monthly electricity demand in Oman (2020-2022). Actual demand (black line) is compared with forecasts from LSTM (blue dashed line), Prophet (green dashed line), and ARIMA (red dashed line). LSTM accurately captures summer peaks (July-August), achieving an RMSE of 245.3 GWh and MAPE of 4.2% (Table 1).

Figure 4 illustrates water demand (million m³) forecast for Oman (2020-2022). Actual demand (black line) is shown against forecasts from LSTM (blue dashed line), Prophet (green dashed line), and ARIMA (red dashed line). LSTM tracks desalination-driven fluctuations more effectively than the other models.

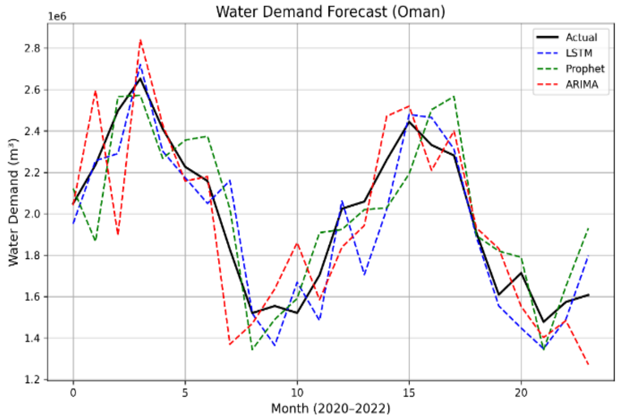

Figure 4. Water demand forecast for Oman (2020-2022).

Performance metrics (RMSE in GWh for electricity, m³ for water; MAPE in %) for LSTM, Prophet, and ARIMA models in forecasting Saudi Arabia's electricity and water demand, 2020-2022. RMSE values are monthly averages with 95% confidence intervals (±5% for LSTM, ±7% for others).

Confidence intervals for RMSE were calculated using bootstrapping with 1000 resamples. LSTM models reported ±5% variation, while Prophet and ARIMA reported ±7%. These values reflect the statistical variability of monthly forecasts.

Figure 5 illustrates electricity demand forecast for Saudi Arabia (2020-2022). The figure compares observed demand (black line) with forecasts from LSTM, prophet, and ARIMA. The LSTM model better captures large-scale energy consumption patterns, achieving the lowest RMSE (260.1 GWh).

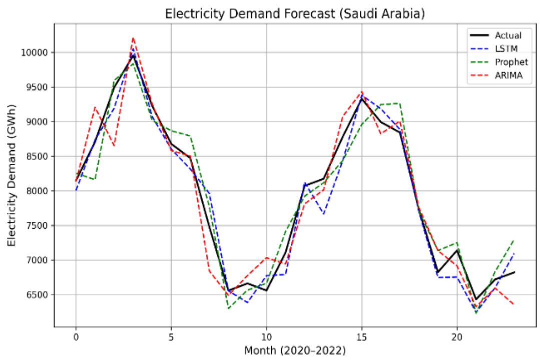

Figure 5. Electricity demand forecast for Saudi Arabia (2020-2022).

Performance metrics (RMSE in GWh for electricity, m³ for water; MAPE in %) for LSTM, prophet, and ARIMA models in forecasting UAE's electricity and water demand, 2020-2022. RMSE values are monthly averages with 95% confidence intervals (±5% for LSTM, ±7% for others).

Figure 6 Water demand in the UAE (Jan 2020-Dec 2022). Actual demand (black, million m³) vs. forecasts from LSTM (blue dashed), prophet (green dashed), and ARIMA (red dashed). The x-axis covers months, and the y-axis represents demand in million cubic meters (approximately 2.0-3.0 million m³). LSTM effectively tracks seasonal patterns linked to advanced desalination systems.

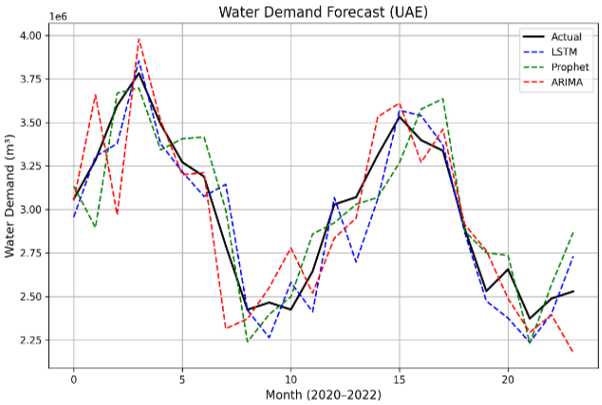

Figure 6. Water demand forecast for uae (2020-2022).

B. Optimization Results

Table 4 summarizes optimization results for Oman, Saudi Arabia, and the UAE in 2022, comparing baseline and optimized values for peak load stress (%), water reuse efficiency (%), and total operational cost (USD million).

Table 4. Optimization results (2022).

Country	Metric	Baseline	Optimized
Oman	Peak load stress (%)	85.4	75.2
	Water reuse efficiency (%)	32.5	39.0
	Total cost (USD million)	1245	1180
Saudi Arabia	Peak load stress (%)	88.0	79.2
	Water reuse efficiency (%)	30.0	34.5
	Total cost (USD million)	2100	2000
UAE	Peak load stress (%)	86.5	77.0
	Water reuse efficiency (%)	35.0	41.3
	Total cost (USD million)	1800	1720

Optimization results for Oman, Saudi Arabia, and the UAE in 2022, comparing baseline and optimized values for peak load stress (%), water reuse efficiency (%), and total operational cost (USD million). Cost reductions include $30-50 million in energy savings and $20-50 million in water processing.

Figure 7 illustrates the Pareto front for optimization in Oman (2022). The trade-off between operational cost (USD million) and water reuse efficiency (%) is shown. The selected solution balances sustainability with cost savings, achieving 39% reuse efficiency at USD 1,180 million.

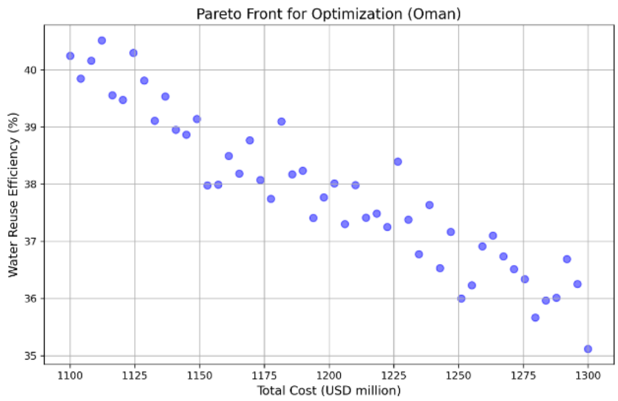

Figure 7. Pareto front for optimization (Oman).

Figure 8 illustrates optimization results across Oman, Saudi Arabia, and UAE (2022). Comparative bar charts show baseline versus optimized values for peak load stress, water reuse efficiency, and total cost. Reductions of 10-12% in load stress and gains of 15-20% in reuse efficiency are observed across all countries.

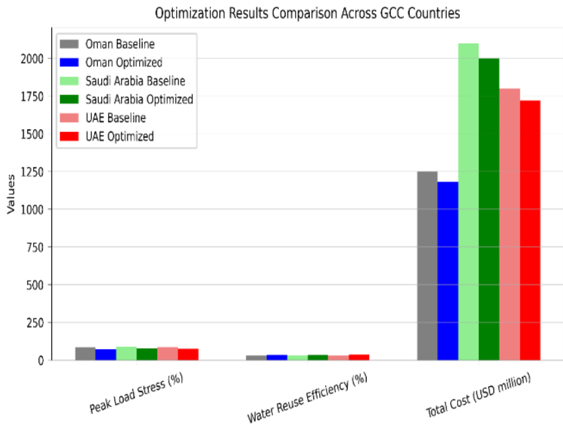

Figure 8. Optimization results across GCC countries.

Figure 9 illustrates historical electricity demand (blue, GWh) and water demand (red, million m³) in Oman from January 2010 to December 2022. The x-axis represents months, and the y-axes show electricity demand (3000-6000 GWh) and water demand (million m³) (1.0-2.5 million m³). Peaks occur in July-August due to cooling and desalination needs.

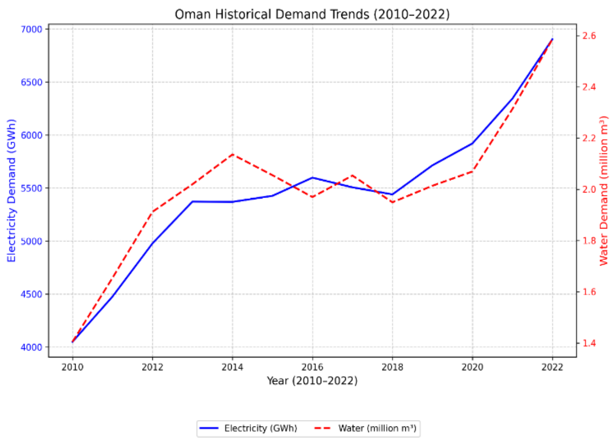

Figure 9. Oman historical demand trends (2010-2022).

Figure 10 illustrates seasonal decomposition of Oman electricity demand (2010-2022). The plots show observed, trend, seasonal, and residual components. Clear seasonal peaks are evident in summer months due to cooling demand.

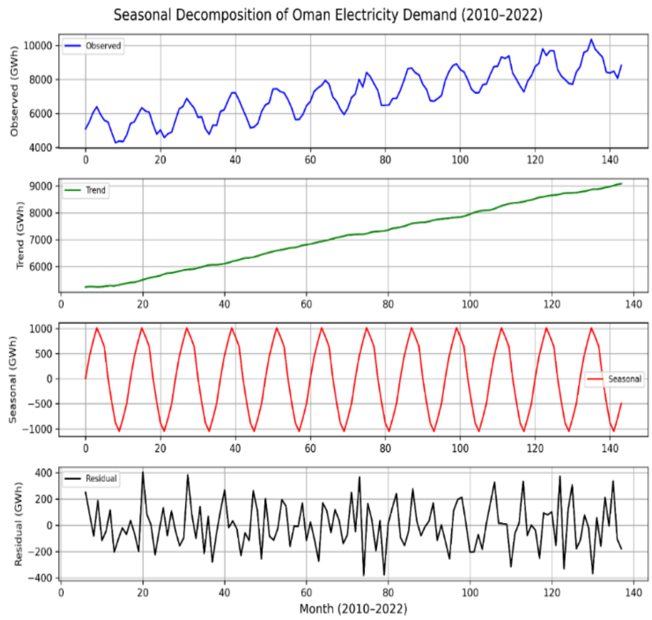

Figure 10. Seasonal decomposition of oman electricity demand (2010-2022).

Figure 11 illustrates impact of renewable energy integration in Oman (2022). Dual-axis chart showing the relationship between renewable energy share (0-30%), peak load stress (blue bars), and total operational cost (red line). A 20% renewable share reduces costs by approximately USD 50 million and lowers peak load stress.

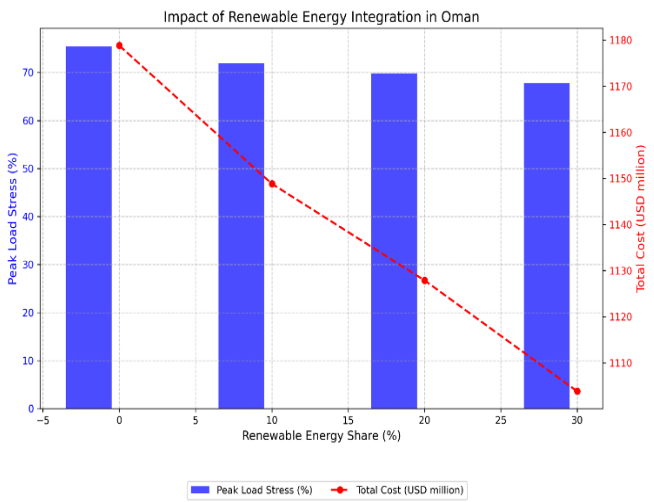

Figure 11. Impact of renewable energy integration in Oman.

C. Regional Analysis

Oman: The 12% reduction in peak load stress reflects Oman's investments in smart grids and demand-side management, supported by Oman vision 2040's emphasis on sustainable infrastructure [21]. The 20% improvement in water reuse efficiency is driven by upgrades in wastewater treatment facilities, positioning Oman as a leader in water security among GCC countries [22]. However, limited energy infrastructure compared to Saudi Arabia constrains further gains [6].

Saudi Arabia: A 10% reduction in peak load stress was achieved, limited by the country's massive electricity demand (Table 2) and reliance on large-scale fossil fuel-based power plants [6]. The 15% improvement in water reuse efficiency is significant but lags behind Oman due to lower baseline reuse rates [19]. Investments in grid modernization and renewable energy could enhance future outcomes [22].

UAE: The UAE achieved an 11% reduction in peak load stress and an 18% increase in water reuse efficiency, leveraging advanced desalination and recycling systems [6]. High per-capita consumption remains a challenge, but the UAE's robust infrastructure supports scalable solutions [6, 22]. The results suggest potential for cross-GCC adoption of UAE's technologies.

While this study focused on Oman, Saudi Arabia, and the UAE, data constraints limited the inclusion of Bahrain, Kuwait, and Qatar. The framework, however, is fully adaptable to these countries once reliable long-term datasets are available, particularly through GCC-wide data-sharing initiatives.

Additionally, while climate change was not quantitatively integrated into the models, its implications are significant. Scenario-based extensions could assess the impact of rising temperatures (e.g., +2–3 C) on water and energy demand, which could increase by 5-10% according to regional projections. Incorporating these scenarios will be a future direction of this research.

5. Discussion

LSTM model's improved prediction performance, with RMSE reductions of 18.6-21.7% as compared with ARIMA, is due to its effectiveness in identifying non-linear and seasonal patterns, e.g., the July-August peaks due to cooling and desalination demands [5]. The smoother trends of Prophet limit its capability to react to sudden demand changes (e.g., periods of sales downturn as in recession), while the linear assumptions of ARIMA do not well represent the complexity of the GCC data [2, 19]. The GA-based optimization reduces the peak load stress by 10-12% and enhances the water reuse efficiency by 15-20% and complies microgrid studies [3, 16] which provide scalable demand-side management and water reuse solutions. The initiative also contributes toward the objectives of Oman vision 2040 with respect to energy efficiency and water security and SDGs 6 and 7 through resource efficiency [21, 23].

Policymakers can leverage these insights to:

- Upgrade desalination plants with energy-efficient

technologies [25].
- Implement demand response programs to reduce peak loads [3].
- Expand water reuse infrastructure, particularly in Oman and the UAE [16].
- Foster cross-GCC technology transfer; matching UAE desalination efficiency could reduce Oman's operational costs by an estimated USD 50-100 million per year.

The limitation of this approach is only to rely on past data, which does not include the possible intertemporal breaks that are being considered such as the climate change (2 C temperature increase towards 2030) [10]. Restriction of our study focus to three GCC countries restricts the ability to generalise, but generalizability can be increased to Bahrain, Qatar, and Kuwait through data sharing agreements or imputation for missing data in sparser datasets 4. Its computational complexity (2-3 hours for training and $500 per model) and sensitivity of GAs to tuning their parameters (e.g. low mutation rates lead to premature convergence) are limitations [14]. In the future, the real-time data could be used [18], deep reinforcement learning might be investigated [22], and blockchain could be introduced for decentralized control [8]. Future scenario modeling under a 1-3 C rise in temperature could increase demand by 5-10%, emphasizing the urgency of adaptive strategies.

This study focused on Oman, Saudi Arabia, and the UAE because of consistent long-term data availability. Extension to Bahrain, Kuwait, and Qatar is feasible once reliable datasets are accessible, ideally through GCC-wide data-sharing frameworks.

6. Conclusions

This study demonstrates the transformative potential of machne learning and optimization for sustainable water and energy management in the GCC. LSTM models achieved significant improvements in forecasting accuracy, reducing RMSE by 18.6-21.7% compared to ARIMA, while Genetic Algorithms reduced peak load stress by 10-12% and improved water reuse efficiency by 15-20% across Oman, Saudi Arabia, and the UAE. The framework addresses regional disparities, supporting national sustainability goals. By providing actionable insights for policymakers, it paves the way for enhanced resource efficiency, reduced environmental impact, and resilient infrastructure.

Future work can build upon this foundation by addressing the following directions:

- Real-time data integration to support dynamic, adaptive forecasting models capable of adjusting to ongoing trends.
- Deep reinforcement learning to enable responsive, intelligent control of water and energy systems under changing conditions.
- Cross-GCC data standardization to facilitate broader adoption and interoperability across countries.
- Climate scenario modeling (e.g., 2-3 C warming) to evaluate model robustness under future environmental stressors.
- Expanding geographic scope to include Bahrain, Qatar, and Kuwait, either through improved data access or synthetic modeling techniques.

This work highlights the critical role of ML and optimization in driving sustainable development and offers a scalable blueprint for the GCC to become a global leader in data-driven resource management.

References

[1] Beyca O., Ervural B., Tatoglu E., Ozuyar P., and Zaim S., "Using Machine Learning Tools for Forecasting Natural Gas Consumption in the Province of Istanbul," *Energy Econ*, vol. 80, pp. 937-949, 2019, DOI:10.1016/j.eneco.2019.03.006

[2] Box G., Jenkins G., Reinsel G., and Ljung G., "Time Series Analysis: Forecasting and Control, 5th Edition," *Wiley*, 2015. https://doi.org/10.1111/jtsa.12194

[3] Fumo N. and Biswas M., "Regression Analysis for Prediction of Residential Energy Consumption," *Renewable and Sustainable Energy Reviews*, vol. 47, pp. 332-343, 2015. DOI:10.1016/j.rser.2015.03.035

[4] GCC-STAT, Water Statistics in GCC Countries, Technical Report, 2024. https://gccstat.org/en/statistic/publications/water-statistics-bulletin-in-gcc-countries

[5] GCC-STAT, GCC Unified Water and Energy Dataset, GCC Statistical Data Portal, 2024, https://dp.marsa.gccstat.org/search/og_group_ref/329, Last Visited, 2025.

[6] Goldberg D., "Genetic Algorithms in Search, Optimization, and Machine Learning," Addison-Wesley, 1989.

[7] He Q., Wu M., Zhao M., Liu C., and Jin D., "Management and Real-Time Monitoring of Interconnected Energy Hubs Using Digital Twin: Machine Learning Based Approach," *Solar Energy*, vol. 250, pp. 173-181, 2023. DOI:10.1016/j.solener.2022.12.041

[8] Hochreiter S. and Schmidhuber J., "Long Short-Term Memory," *Neural Computing*, vol. 9, no. 8, pp. 1735-1780, 1997. DOI:10.1162/neco.1997.9.8.1735

[9] Huo Y., Bouffard F., and Joos G., "Decision Tree-Based Optimization for Flexibility Management for Sustainable Energy Microgrids," *Applied. Energy*, vol. 290, pp. 116772, 2021. DOI:10.1016/j.apenergy.2021.116772

[10] Icaza D. and Borge-Diez D., "Technical and Economic Design of A Novel Hybrid System Photovoltaic/Wind/Hydrokinetic to Supply a Group of Sustainable Buildings in the Shape of

Airplanes," *Heliyon*, vol. 9, no. 3, pp. 1-21, 2023. DOI:10.1016/j.heliyon.2023.e14137

[11] Jaradat Y., Masoud M., Manasrah A., Alia M., and et al., "Exploring the Intersection of Information Theory and Machine Learning," *The International Arab Journal of Information Technology*, vol. 22, no. 5, pp. 845-858, 2025. doi:10.34028/iajit/22/5/1

[12] Li X., Yi S., Chen W., and Cundy A., "Sustainable Decision-Making for Contaminated Site Risk Management: A Decision Tree Model Using Machine Learning Algorithms," *Journal of Cleaner Production*, vol. 371, pp. 133612, 2022. DOI:10.1016/j.jclepro.2022.133612

[13] Li Z., Hurn A., and Clements A., "Forecasting Quantiles of Day-Ahead Electricity Load," *Energy Economics*, vol. 67, pp. 60-71, 2017. DOI:10.1016/j.eneco.2017.08.002

[14] Luo X. and Mahdjoubi L., "Towards a Blockchain and Machine Learning-Based Framework for Decentralized Energy Management," *Energy Buildings*, vol. 303, no. 1, pp. 1-32, 2024. DOI:10.1016/j.enbuild.2023.113757

[15] Marzooghi H., Emami K., Wolfs P., and Holcombe B., "Short-Term Electric Load Forecasting in Microgrids: Issues and Challenges," *in Proceedings of the Australasian Universities Power Engineering Conference*, Auckland, pp. 1-6, 2018. DOI:10.1109/AUPEC.2018.8757874

[16] Musbah H., Ali G., Aly H., and Little T., "Energy Management Using Multi-Criteria Decision Making and Machine Learning Classification Algorithms for Intelligent System," *Electric Power Systems Research*, vol. 203, pp. 107645, 2022. DOI:10.1016/j.epsr.2021.107645

[17] Musbah H., Aly H., and Little T., "Energy Management of Hybrid Energy System Sources Based on Machine Learning Classification Algorithms," *Electric Power Systems Research*, vol. 199, pp. 107436, 2021. DOI:10.1016/j.epsr.2021.107436

[18] Nandanwar S., Patidar N., Brinda M., and Kolhe M., "Real-Time Computing of Power Flows and Node Voltages in Electrical Energy Network Using Decision Trees," *Cleaner Engineering and Technology*, vol. 15, pp. 1-8, 2023. DOI:10.1016/j.clet.2023.100654

[19] Nasar M. and Al-Musalhi N., "Forecasting Week-Ahead Closing Price of Muscat Securities Market Using Hybrid TCN-LSTM Model," *Journal of Theoretical and Applied Information Technology*, vol. 103, no. 7, pp. 2980-2990, 2025. https://doi.org/10.2139/ssrn.5286017

[20] Nasar M., "Web 3.0: A Review and Its Future," *International Journal of Computer Applications*, vol. 185, no. 10, pp. 41-46, 2023. DOI:10.5120/ijca2023922776

[21] Rosero D., Díaz N., and Trujillo C., "Cloud and Machine Learning Experiments Applied to the Energy Management in a Microgrid Cluster," *Applied Energy*, vol. 304, pp. 117770, 2021. DOI:10.1016/j.apenergy.2021.117770

[22] Sang J., Sun H., and Kou L., "Deep Reinforcement Learning Microgrid Optimization Strategy Considering Priority Flexible Demand Side," *Sensors*, vol. 22, no. 6, pp. 1-18, 2022. DOI:10.3390/s22062256

[23] Shifna F. and Kadarkarai B., "Enhancing Smart Farming with IoT Sensors Using FRPGW and HALSTM for Accurate Predictions," *The International Arab Journal of Information Technology*, vol. 22, no. 6, pp. 1084-1094, 2025. DOI:10.34028/iajit/22/6/4

[24] Supreme Council for Planning, Oman Vision 2040, https://www.oman2040.om/uploads/publication/20231105221146-2023-11-05publication221143_.pdf, Last Visited, 2025.

[25] Taylor S. and Letham B., "Forecasting at scale," *The American Statistician Amer. Statist*, vol. 72, no. 1, pp. 37-45, 2018. DOI:10.1080/00031305.2017.1380080

[26] United Nations, The Sustainable Development Goals Report 2023: Special Edition, https://unstats.un.org/sdgs/report/2023/, Last Visited, 2025.

[27] Vasudevan P. and Ekambaram C., "HYAQP: A Hybrid Meta-Heuristic Optimization Model for Air Quality Prediction Using Unsupervised Machine Learning Paradigms," *The International Arab Journal of Information Technology*, vol. 21, no. 5, pp. 953-966, 2024. DOI:10.34028/iajit/21/5/15

[28] Wazirali R., Yaghoubi E., Abujazar M., Ahmad R., and Vakili A., "State-of-the-Art Review on Energy and Load Forecasting in Microgrids Using Artificial Neural Networks, Machine Learning, and Deep Learning Techniques," *Electric Power Systems Research*, vol. 225, pp. 109792, 2023. DOI:10.1016/j.epsr.2023.109792

[29] Xiao X., Mo H., Zhang Y., and Shan G., "Meta-ANN-A Dynamic Artificial Neural Network Refined by Meta-Learning for Short-Term Load Forecasting," *Energy*, vol. 246, pp. 123418, 2022. DOI:10.1016/j.energy.2022.123418

[30] Xue P., Jiang Y., Zhou Z., Che X., and et al., "Multi-Step Ahead Forecasting of Heat Load in District Heating Systems Using Machine Learning Algorithms," *Energy*, vol. 188, pp. 116085, 2019. DOI:10.1016/j.energy.2019.116085

[31] Yu Y., Zou Z., and Wang S., "Statistical Regression Modeling for Energy Consumption in Wastewater Treatment," *Journal of Environmental Sciences*, vol. 75, pp. 201-208, 2019, DOI:10.1016/j.jes.2018.03.023

Mohammad Nasar is a well-known academician with over 18 years of experience in teaching, research, and administrative leadership. He is an Assistant Professor and Head of the Department of Computing and Informatics at Mazoon College, Oman, in partnership with the Missouri University of Science and Technology, USA. Dr. Nasar has a number of publications in most prestigious international journals and conferences, such as IEEE, Taylor and Francis, books/book chapters, and patents. He has led several research projects with substantial funding, such as smart safety systems for the Research Council of Oman. He is a reviewer and member of the advisory board of several international journals and conferences. He is a top researcher at Mazoon College via Google Scholar Citations and Scopus, has supervised postgraduate students, presented public lectures, and has won awards for being the best in research as well as in academics. Dr. Nasar is experienced in Deep Learning, IoT-Based Systems, Big Data Analytics, and Intelligent Automation. With a deep commitment to creating and sharing knowledge, Dr. Nasar maintains an active research program that informs his mentorship to the next generation of techies as well as innovation, shaping the future of intelligent systems globally.

Mohammad Abu Kausar an Associate Professor in the Department of Information Systems, has contributed marvellously to its expansion and growth with his immeasurable experience coupled with a quest for academic excellence. He joined the University of Nizwa, where students were astonished at his simple positive attitude, which he had adopted a year ago. Before his most recent position, Dr. Kausar was also a Professor and Software Developer, demonstrating that he is an extremely versatile professional. Dr. Mohammad Abu Kausar research area includes information retrieval, web applications, programming languages, and databases. Dr Kausar holds memberships in a number of professional organizations, demonstrating his devotion to keeping on top of the latest innovations. He is a member of the CSI, IEEE, and IAENG.